Mobvista 移动广告

反作弊白皮书

Mobvista

目录

前言

目前全球移动广告和变现市场正处在高速发展期。App Annie的数据显示，2019年移动应用的下载量达到了2040亿次，预计到2022年将达到2580亿次。全球市场的移动广告支出也在逐年上升，根据市场研究机构eMarketer的数据显示，2019年移动广告支出超过了2200亿美元，到2022年预计将攀升到3521亿美元。

正所谓有利益的地方就有犯罪，日益猖獗的移动广告作弊也深深地困扰着整个广告行业。根据WFA的保守估计，2019年移动广告作弊行为造成广告主的损失超过200亿美元，并预计到2025年，该数字将达到500亿美元之巨。移动广告作弊不仅仅给广告主造成了巨大的经济损失，同时也影响了广告平台的交付质量和品牌形象，更是严重损害了整个移动互联网广告市场长期的良性发展。

Mobvista定位为向全球移动开发者提供用户增长和流量变现的综合性服务的技术平台，建立的广告网络覆盖超过200+个国家和地区，每天处理请求数超过100亿次，可以说是处于反作弊攻防这场战争的最前线。

为净化互联网广告市场、促进行业的健康发展，Mobvista向行业推出《移动广告反作弊白皮书2.0》，旨在厘清当前移动营销市场的广告作弊情况、作弊方式及反作弊策略，以期推动移动广告行业的透明度和规范化发展。

什么是移动广告作弊

移动广告作弊（Mobile Ad Fraud），是指通过技术手段欺骗广告主、发行商或供应商的行为。作弊的主要目的是窃取广告主的广告预算。

移动广告作弊有多种形式，常见的有Click Spamming、Click Injection、SDK Spoofing、Bots等。本章节我们会详细地讨论常见的作弊方式。

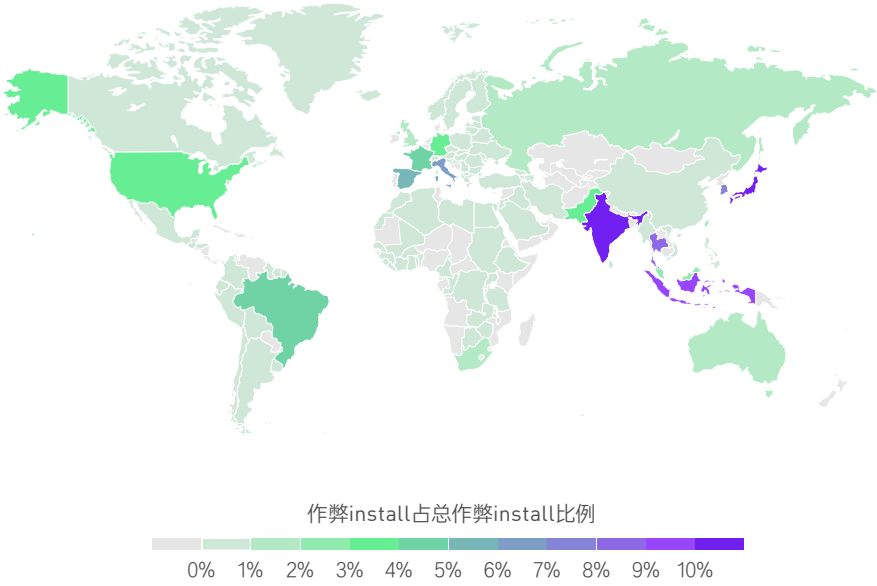
作弊现状

基于Mobvista旗下的程序化互动式移动广告平台Mintegral和移动效果营销平台Nativex覆盖200+个国家（地区）的流量体系，100亿日展示等自有数据，Mobvista反作弊技术团队对全球移动作弊安装行为进行了统计分析（2019年10月-2020年3月），Mintegral被监测并提前拦截的作弊流量占总流量的比例为2.01%，Nativex被监测并提前拦截的作弊流量占总流量的比例为10.69%。安装作弊行为在地区、操作系统、APP类型上的分布呈现以下特点：

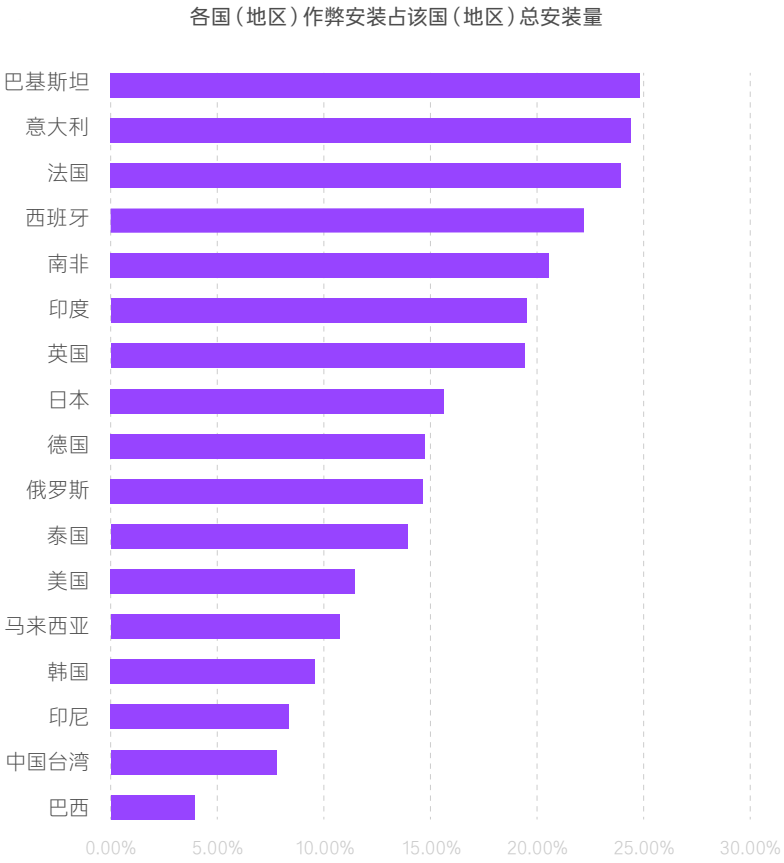
地区分布

在Mintegral和Nativex覆盖200+个国家和地区的流量体系中，被拦截的作弊行为主要分布在日本（12.81%），印度（11.6%），印尼（9.4%）等国家。（注：由于以上国家（地区）的流量在Mintegral和Nativex流量体系中的占比本身较高，该国家（地区）产生的作弊安装与总作弊安装之比自然较大，因此参考意义有限。）

作弊国家（地区）分布图



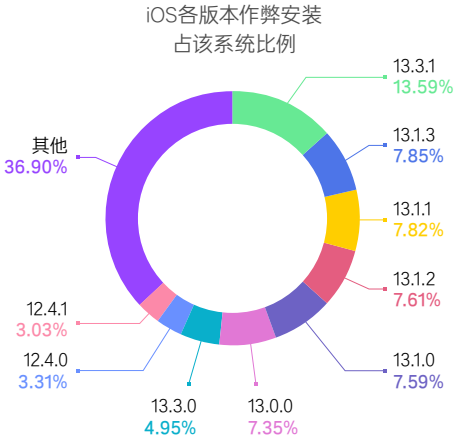
为了能更直观地看到各国家（地区）作弊行为的严重程度，这里同时统计了各国家（地区）作弊安装占该国家（地区）总安装量的比例，比较严重的有：巴基斯坦（24.87%），意大利（24.47%），法国（23.97%）：

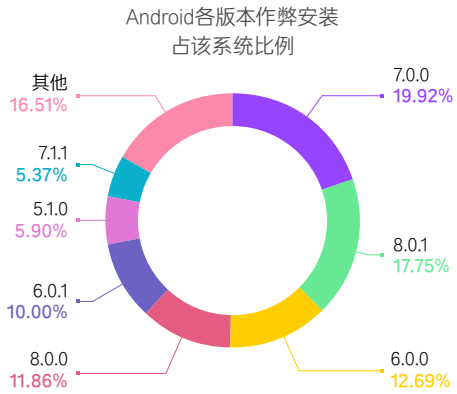


操作系统分布

从操作系统的作弊安装分布来看，iOS平台为40.2%，Android平台为59.8%。

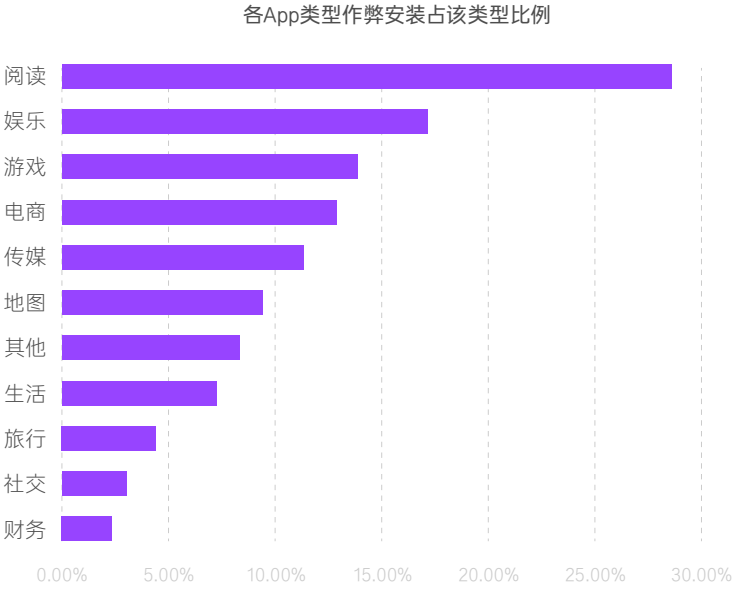
其中，iOS的作弊安装主要集中在这几个版本上：13.3.1(13.59%)，13.1.3(7.85%)，13.1.1(7.82%)。Android平台的作弊安装主要集中在这几个版本上：7.0.0(19.92%)，8.1.0(17.75%)，6.0.0(12.69%)。





APP类型分布

根据对游戏、电商、社交、工具等类型的APP进行的统计数据，不同APP类型作弊安装占该APP类型总安装的比例不尽相同，其中作弊比例较高的有阅读（28.60%），娱乐（17.17%），游戏（13.86%）。

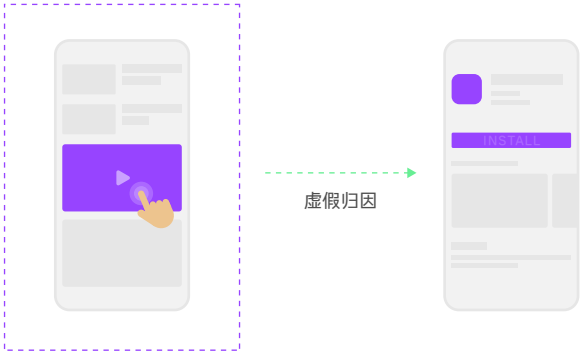


常见的作弊行为

根据流量的真实性，可将常见的作弊行为分为归因作弊、虚假流量和非法流量三大类。

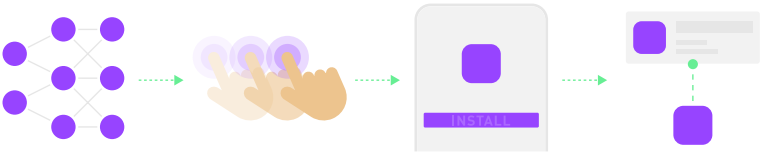
归因作弊

归因作弊，是指作弊者利用第三方归因策略（多为最近展现/点击归因）的漏洞，虚构展现或点击以窃取用户的安装的作弊行为。



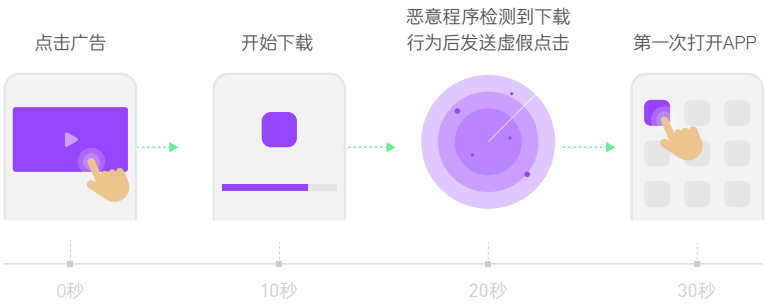
A. Click Spamming

Click Spamming，也叫Click Stuffing或Click Flood，是指作弊者利用点击归因的漏洞，通过发送大量的虚假点击来窃取自然用户的安装，从而造成广告主预算的浪费。这种作弊流量的特点是转化率(CVR)偏低，点击到转化时间普遍过长。



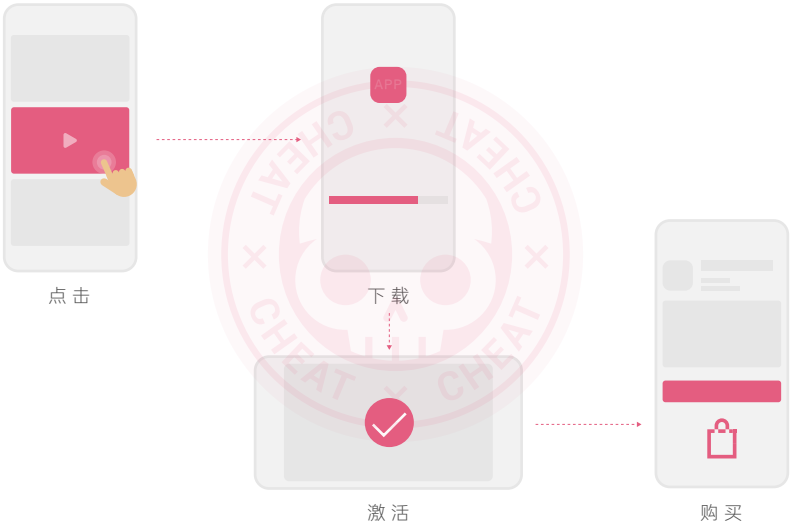
B. Click Injection

Click Injection，也叫Install Hijacking，指的是作弊者通过安装在用户设备上的一个应用程序来“监听”其他应用程序的安装广播消息。当用户设备上安装了新的应用程序时，作弊者就会收到通知，然后在安装完成之前发送虚假点击利用归因模型的漏洞劫取相应的安装。特点是点击到安装时间过短，应用商店记录的下载时间早于点击广告的时间。



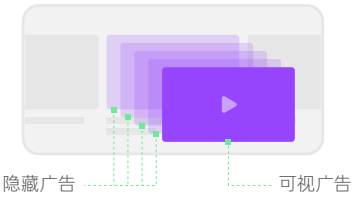
虚假流量

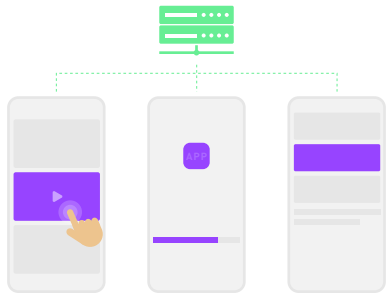
虚假流量，指的是广告过程中的点击、安装、应用内行为是无效的，广告主花了钱却不能获得一个真实用户。



A. AD Stacking

当用户点击一个广告时，publisher 却向后台发送多个不同广告的点击，从而吃掉大量CPC的广告预算。特点是同一个设备短时间内多次点击广告。



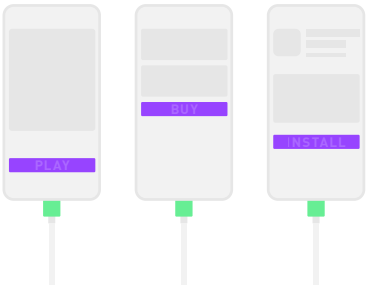


B. Bots

Bots 指的是作弊者通过自动化脚本或计算机程序模拟真实用户的点击、下载、安装甚至是应用内行为，伪装成为真实用户，从而骗取广告主的CPI/CPA预算。特点是IP离散度密集、新设备率过高、用户行为异常、机型/系统/时间等分布异常等。

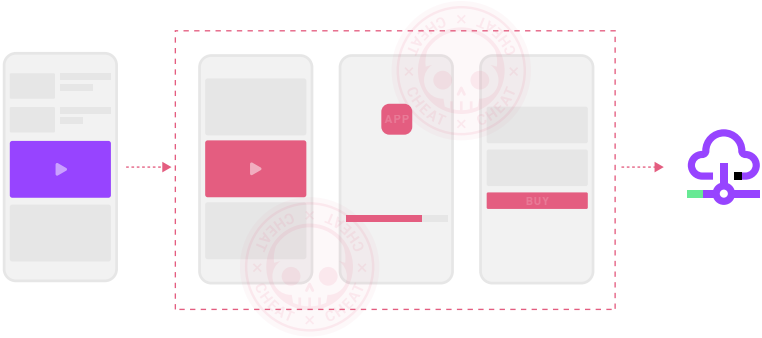
C. Device Farms

Device Farms指的是作弊者购买大量真实设备进行广告点击、下载、安装和应用内行为，并通过修改设备广告跟踪符等方式隐藏设备信息。特点是IP离散度密集、新设备率过高、用户行为异常、机型/系统分布异常等。



D. SDK Spoofing

SDK Spoofing是指作弊者通过执行“中间人攻击”破解第三方SDK的通信协议后，在没有任何实际安装的情况下，使用真实设备的数据来发送虚假的点击和安装，以此消耗广告主的预算的作弊行为。特点是广告主后台数据和第三方数据不符。



非法流量

非法流量指的是publisher在未经广告主同意的情况下，使用一些有争议的手段来获取用户，包括从非官方应用商店下载app、激励流量、被禁止的广告素材、网赚、诱骗点击和下载、木马后台操作等。

如何反作弊

反作弊行业介绍

移动广告反作弊业内的主要角色包括：

a. 广告主

b. 第三方检测平台，包括德国的Adjust、以色列的AppsFlyer、美国的TUNE、Kochava、国内的TalkingData等。

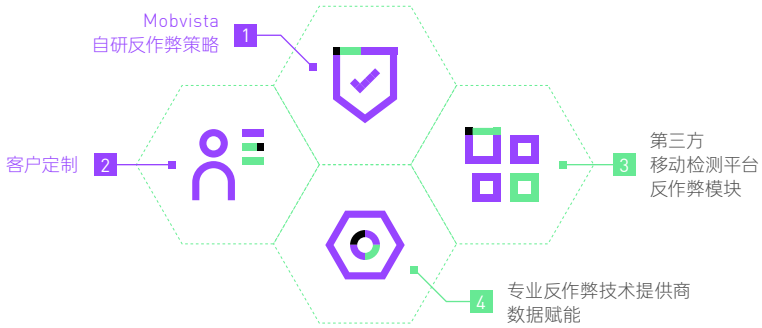
c. 广告平台

d. 其他反作弊技术服务提供商，常见的有英国的Machine、印度的mFilterIt、美国的FraudScore和DataVisor等。

Mobvista 反作弊体系

Mobvista 反作弊体系简介

Mobvista反作弊技术体系，是以自研的反作弊技术为主，通过融汇、协同第三方移动检测平台、其他反作弊技术服务提供商的反作弊服务为辅，构建的多层次反作弊技术体系。



Mobvista的反作弊力量主要由四部分构成。

其中，客户定制部分，主要包括根据广告主的event数据回传，进行关键指标及阈值的确认、定制数据监控等，即Mobvista根据广告主的具体需求制定反作弊策略。

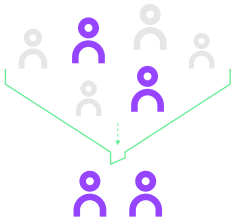
第三方移动检测平台反作弊策略，指Mobvista参考并纳入适用的第三方平台的反作弊策略，如基于Adjust的IP黑名单完善Mobvista的IP过滤机制。

专业反作弊技术提供商数据赋能，指借助专业反作弊技术公司的反作弊服务进行数据分析、多重监测、筛选流量，最大程度减少作弊带来的损失。目前Mobvista合作的该领域公司包括美国的Fraudlogix，马来西亚的ip2location。

而最为核心的Mobvista的自研反作弊技术系统，它是基于Mobvista积累的作弊特征数据建立的，以规则为主、机器学习模型为辅的方式，实现了对作弊流量的多层次全方位的识别。

基于规则的反作弊策略

目前，Mobvista的反作弊技术团队已积累60+的特征，并时刻紧跟业界最新研究不断增加新的特征，以此为基础建立了完善的反作弊规则体系，以下呈现部分反作弊规则。



IP相关

A. 来自相同IP的点击/安装量过多

正常情况下，由于用户分布的离散性，用户设备使用网络的IP分布也应该呈现出离散性。如果点击或安装数据异常集中地来自相同IP，那么有一定可能是作弊行为。（需要注意的是，由于不同运营商的IP池大小不同，IP分配策略也不同，以及存在同一局域网内的多台设备同时连接到公网的可能，只要不过度集中在某几个IP上，一定的IP重复都是正常的。）

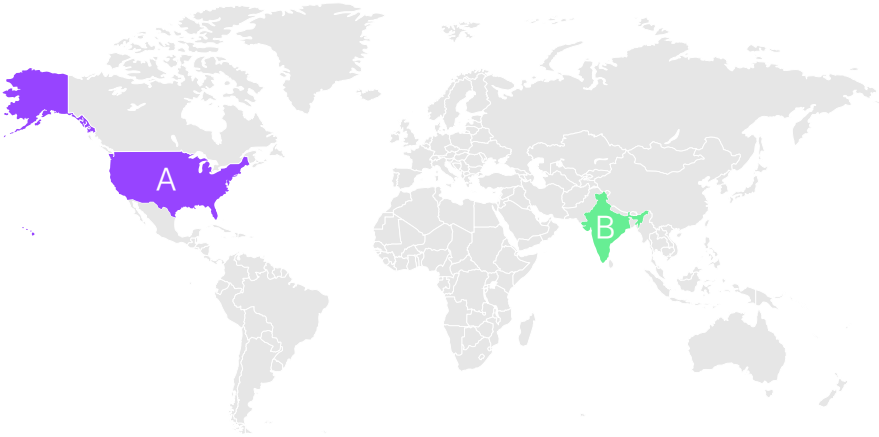
B. IP黑名单

IP黑名单主要来自于第三方反作弊服务商，和Mobvista长期以来的积累数据，一般主要是不可能成为移动设备IP的地址（如数据中心等IP），或者有过大量作弊历史的高风险IP。

C. 点击IP和安装IP不一致比例过高

大部分情况下，点击IP和安装IP应该是相同的。如果一个渠道出现点击IP和安装IP不一致比例过高的情况，那么这个渠道就存在一定的作弊风险。

地理相关



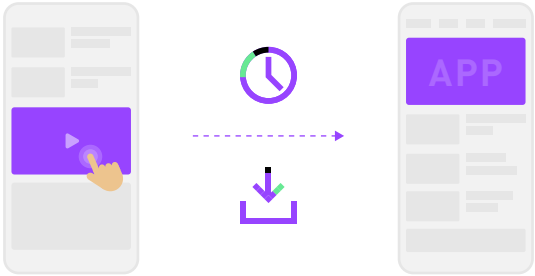
A. 安装国家与投放国家不一致

如果广告投放的目的地是美国，而最终完成安装的IP在印度，此类场景在现实中出现的概率很低，如果大规模出现，则有很高的作弊可能性。

B. 安装城市与点击城市不一致

如果点击IP在北京，而完成安装的IP在广州，此类场景的概率同样也很低，如果大规模出现，则有很高的作弊可能性。

点击到完成安装的时间



A. 点击到完成安装的时间过短

通常用户从点击广告、跳转到应用市场App的下载页面、下载App、到最终打开App进行激活，需要一定的时间。如果该时间过短，则很有可能是作弊行为；这个时间具体多少才算过短，需要根据当地的网速、app安装包的大小、手机性能而定。

B. 点击到完成安装的时间过长

单独的一次安装过程，点击到完成安装的时间过长可能是偶然事件，但如果这样的场景大比例出现（例如80%的点击到完成安装的时间大于24小时），则有很大可能性是作弊行为。

设备标记符

设备标记符是用来标识一台移动设备的ID，被开发者或广告主用来追踪广告效果，对于iOS设备通常采用IDFA，Android设备通常采用GAID。由于具备足够的随机性，实际应用中冲突很少，基本上可以认为能标记唯一的一台移动设备。

IDFA | GAID

A. 新设备率过高

为了保证隐私，设备标记符可以被手动修改，但大部分用户一般不会主动修改设备标记符，所以全新的标记符一般要么出现在此前未通过广告安装App的设备上，要么出现在试图通过频繁修改标记符安装App的作弊设备上。基于此，如果一个渠道的新设备比例过高，那么肯定会有很大的作弊嫌疑。

B. 设备黑名单

通过大量历史积累数据，Mobvista会把一些有频繁作弊历史的设备ID列入黑名单，如果一个安装发生在黑名单的设备ID上，那么该安装也存在很高的风险。

设备参数

渠道设备的品牌、型号、语言、运营商、User Agent，操作系统等设备参数的具体分布情况，必须和当地实际分布相符合。例如，如果一个渠道安装的大部分设备品牌型号，都不是当地的主流型号，则该渠道肯定会有很大的作弊嫌疑。

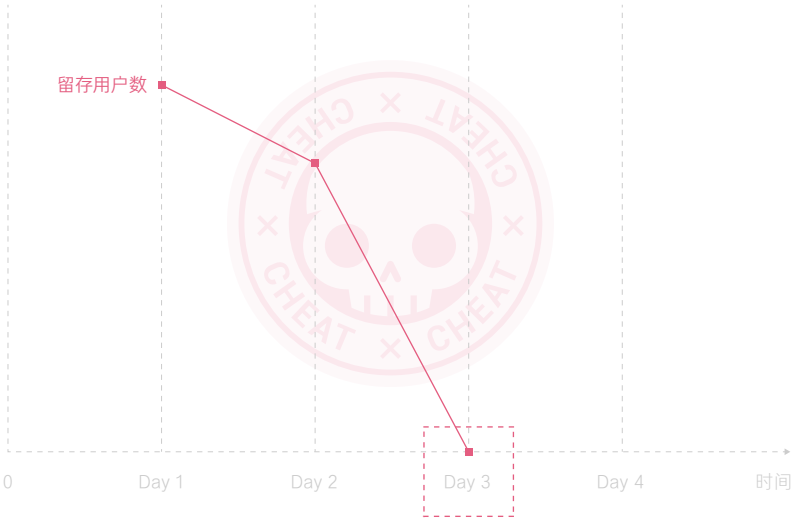
用户行为

A. 极端低/高的转化率

正常情况下渠道的转化率会在一定的区间范围内，转化率过低或者过高的渠道都需要特别关注。例如，对于转化率过高的渠道需要注意其是否是非法流量。

B. 留存曲线发生明显下滑

如果某一个渠道带来的用户总是在游戏/应用中的某一时间节点，数据表现出断崖式下滑，例如用户安装应用后次日留存正常，但第3天留存降为0，则有很高的作弊嫌疑。



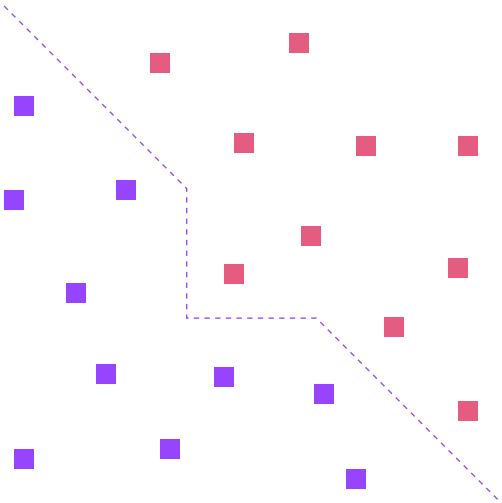
如何解决混合流量

必须注意的是，越来越多的作弊流量并不是单一的作弊类型，而是多种作弊类型相互混合后的结果。比如，将click spamming和bots流量混合在一起，既能在一定程度上提高渠道转化率，同时又能缓解点击到安装转化的时间差普遍过长的问题，这大大增加了反作弊的难度。

要解决这种多作弊类型流量混合的问题，就必须在统计渠道的数据分布和指标时，增加统计的维度，在高维度上将低维度难以分解的流量分解出来。

除了传统的渠道维度、平台维度等，还可用于升维的维度有：设备维度（如品牌、型号、操作系统、运营商等），地理维度（如城市、IP子网、省、州、邦等），时间维度（如点击到安装时间，一天中的某个小时节点等）。

比如作弊者将click spamming流量和bots流量混合在一起，使得原始流量的各项指标实现了“互补”，对识别作弊造成了困难。其中click spamming流量主要发生在三星手机上，bots流量主要发生在华为手机上。这时我们通过增加手机品牌这一统计维度，就可以拆分出原始流量的指标，还原原始流量的真实情况。

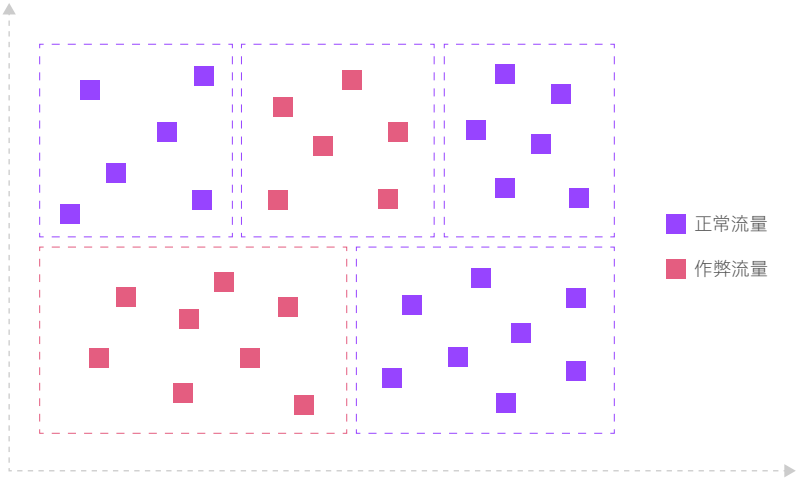


基于机器学习的反作弊策略

通过一系列反作弊规则来阻止特定的作弊模式的作弊流量一直是互联网广告行业的标准做法。虽然基于规则的反作弊解决方案是一种有效直观的解决方案，但是它们却有一定的滞后性，不能灵活适应新的作弊模式，并且随着作弊手段/模式不断进化演变，作弊模式变得更加复杂，从业人员需要采用越来越复杂的方法来识别相应的作弊流量。而基于机器学习的反作弊解决方案能够从海量的数据中自动学习和发现潜在的或者新的作弊模式。

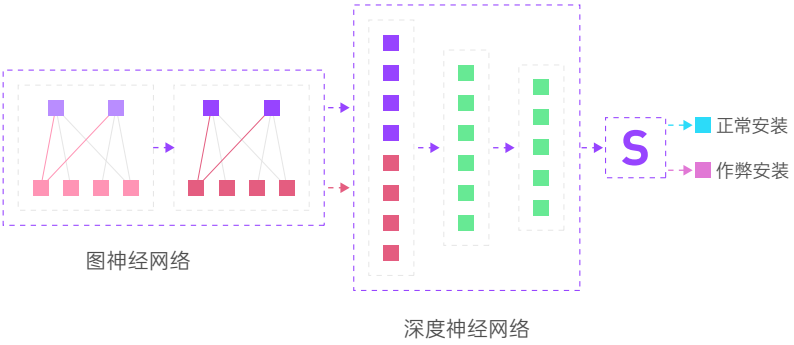
无监督反作弊模型

无监督机器学习模型无需事先对数据进行标注(label)，通过对数据的探索和分组就能够自动发现潜在的或者新的作弊模式。常用的无监督机器学习有聚类、异常检测等。例如在Mobvista我们采用了Density Peak Clustering对渠道/siteid等进行聚类，人工专家根据聚类的结果再做进一步的分析判断相应的流量是否是作弊流量。



有监督反作弊模型

对于有标注的数据，我们可以通过有监督机器学习来建立相应的反作弊模型。反作弊领域常用的有监督模型有Random Forest、GBDT以及近两年兴起的图神经网络(GNN)。例如在Mobvista我们采用了GNN来识别Bots作弊流量，模型的准确率高达92%。



结语

在本白皮书中，我们详细描述了移动广告领域各类常见的作弊类型，并针对性介绍了Mobvista的反作弊技术体系。集团副总裁朱亚东博士表示，作弊和反作弊之间的攻防战是一场持久战，随着作弊手段和技术的不断升级，反作弊技术和措施也在不断演进。为了维护行业的健康发展，我们呼吁广告主、广告代理公司、第三方检测公司等行业组织在和打击作弊流量上需要形成共识，以期推动移动广告行业的透明度和规范化发展。

关于Mobvista

Mobvista作为全球领先的技术平台，致力于推动数字时代的全球商业增长。

凭借全球技术和丰富的行业经验，Mobvista专注于打造工具生态系统(Tooling Ecosystem)，帮助企业利用大数据、人工智能、云计算弹性集群管理等先进技术连接中国与世界，帮助企业构建具有前瞻性的业务模式，以更高的效率和效能触达市场。

Mobvista集团旗下现有三大业务品牌：程序化互动式移动广告平台Mintegral、移动效果营销平台Nativex、移动游戏数据分析平台GameAnalytics。

Mobvista于2013年在中国广州成立，并于2018年12月在香港联合交易所主板上市(01860.HK)。目前，Mobvista拥有700多名员工，在全球16个城市设有办事处。

Mobvista

www.mobvista.com